Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6, 7951-7970
2024
Publisher: Learning Gate
DOI: 10.55214/25768484.v8i6.3732
© 2024 by the authors; licensee Learning Gate

Two self-supervised image denoiser designs with discrete wavelet transform and non-local means-based algorithms

Ashishkumar Gor[1*], C.K. Bhensdadia[2]
[1,2]Computer Engineering, Dharmsinh Desai University, Nadiad, 387001, Gujarat, India; ashishgor.ce@ddu.ac.in (A.G.) ckbhensdadia@ddu.ac.in (C.K.B.).

Abstract: Image denoising is crucial in applications like medical imaging and photography, where restoring high-quality images from noisy data is essential. Traditional techniques often struggle with complex noise patterns, while deep learning-based methods typically rely on clean-noisy image pairs for training, limiting their practicality. Additionally, deep learning approaches face challenges such as the lack of ground truth clean images, sensitivity to specific noise types, and the introduction of artifacts during processing. In this work, we propose two novel self-supervised denoising approaches: a Discrete Wavelet Transform (DWT)-based model and a Non-Local Means (NLM)-based model. The DWT-based approach employs wavelet decomposition to separate image details across multiple frequency scales, selectively suppressing high-frequency noise via soft thresholding while preserving low-frequency components. The resulting wavelet coefficients are used to create pseudo-clean targets for training a U-Net architecture, ensuring effective denoising while maintaining structural integrity. The NLM-based approach leverages redundancy in image patches by applying the NLM algorithm to generate pseudo-clean targets through patch similarity-based averaging. These targets train a U-Net model with a custom loss function that balances Mean Squared Error (MSE), Peak Signal-to-Noise Ratio (PSNR), and Structural Similarity Index (SSIM), optimizing perceptual quality. Both models are trained on 5,000 noisy images from the ImageNet validation set without relying on clean references. Validated on synthetic Gaussian and Poisson noise at varying magnitudes, the DWT-based model achieved Mean PSNR and SSIM values of 31.07 and 0.9279, respectively, while the NLM-based model attained 30.17 and 0.9303. These results demonstrate the robustness and effectiveness of the proposed methods, making them suitable for real-world applications such as medical diagnostics and low-light photography.

Keywords: *Discrete wavelet transforms (DWT), Gaussian noise, Poisson noise, Haar wavelet, Debauchies wavelet, Image denoising, Non-local means (NLM), PSNR (Peak Signal-to-noise ratio), Self-supervised learning, SSIM (Structural similarity index), Thresholding.*

1. Introduction

Image denoising is a fundamental task in computer vision, critical for applications such as medical imaging, photography, and astronomy, where image quality significantly impacts downstream analysis. The objective is to restore images degraded by noise, ensuring the preservation of essential features. Traditional and deep learning-based methods have dominated this field, with recent advancements in self-supervised learning (SSL) offering robust solutions to address the limitations of clean-noisy paired datasets. This section reviews existing approaches while contextualizing the proposed methods based on the Discrete Wavelet Transform (DWT) and Non-Local Means (NLM).

Traditional Denoising Methods: Traditional denoisers rely on mathematical models and handcrafted priors. Block Matching and 3D Filtering (BM3D) [1] groups similar patches through block matching and applies collaborative filtering in the transform domain, effectively reducing noise while retaining image features. Non-Local Means (NLM) [2-4] improves upon local smoothing techniques by averaging pixels based on patch similarity across the entire image, leveraging spatial redundancy to

© 2024 by the authors; licensee Learning Gate
* Correspondence: ashishgor.ce@ddu.ac.in

enhance detail preservation. Weighted Nuclear Norm Minimization (WNNM) [5] adopts low-rank matrix approximations, preserving image structures while suppressing noise. Despite their effectiveness, these methods struggle with high computational costs and adapting to complex noise patterns in real-world scenarios.

Supervised Learning for Image Denoising: The advent of deep learning introduced supervised models like DnCNN [6], which combines convolutional networks with residual learning to map noisy images to clean counterparts. While these methods outperform traditional approaches, they depend on large paired datasets, which are expensive and challenging to acquire in specialized domains. Additionally, their performance is often tailored to specific noise distributions, limiting their robustness to unseen noise types [7].

Emergence of Self-Supervised Learning: Self-supervised learning eliminates the dependency on clean images by training models directly on noisy inputs. Noise2Noise [7], a pioneering approach, demonstrated that noisy images alone could serve as training data by treating one noisy image as the target for another. This concept led to blind-spot methods like Noise2Void [8] and Noise2Self [9], which train models by masking certain pixels or regions during prediction. These methods leverage the intrinsic structure of the image to avoid learning identity mappings, achieving robust denoising without paired data.

Wavelet-Based Denoising: Wavelet transforms, particularly Discrete Wavelet Transform (DWT), provide a multi-resolution framework for analysing images at different frequency scales. DWT decomposes images into approximation and detail coefficients, allowing selective denoising by thresholding high-frequency components associated with noise. Wavelet families like Haar, Daubechies, and Symlets offer flexibility for various denoising tasks. Classical wavelet denoising, though effective in suppressing noise, often introduces artifacts like blurring and ringing [10,11]. Recent advancements integrate DWT with deep learning. Liu and Liu [12] combined DWT with CNNs, denoising wavelet coefficients using learned filters, achieving improved performance compared to traditional CNNs. However, such hybrid approaches often face computational challenges.

Non-Local Means-Based Denoising: NLM extends traditional spatial filtering by averaging pixel values based on the similarity of patches across the entire image. This algorithm is particularly effective for reducing Gaussian noise while preserving edges and textures. While computationally intensive, NLM offers robustness and adaptability, making it a strong candidate for integration into deep learning frameworks. Recent efforts have combined NLM with self-supervised methods, using pseudo-clean images derived from NLM as training target.

Recent Self-Supervised Innovations: Several innovative SSL methods address diverse noise scenarios. Blind2Unblind [13] enhances blind-spot models with global-aware masking, improving training diversity. Recorrupted2Recorrupted [14] extends Noise2Noise by introducing multiple noise levels during training, enabling models to generalize across varied noise conditions. Neighbor2Neighbor [15] creates training pairs by subsampling noisy images, maintaining spatial consistency without requiring explicit noise models. These advancements underscore the flexibility of self-supervised learning in addressing real-world challenges.

Contextualizing the Proposed Methods: The proposed self-supervised denoising methods build on these advancements, leveraging DWT and NLM within an SSL framework. The DWT-based approach uses thresholded wavelet coefficients as pseudo-clean targets, enabling multi-resolution denoising without requiring clean references. By combining DWT with a U-Net architecture [16], this method effectively suppresses noise while preserving structural details, achieving state-of-the-art performance on Gaussian and Poisson noise.

The NLM-based approach generates pseudo-clean images using NLM, which are then used to train a U-Net model [16] with a custom loss function balancing Mean Squared Error (MSE) and Structural Similarity Index (SSIM) [17]. This integration allows the model to leverage both spatial and structural redundancies, enhancing denoising performance across varied noise types. This highlights the evolution of image denoising, from traditional methods to modern self-supervised approaches. While supervised learning has advanced denoising performance, self-supervised methods like Noise2Noise and Neighbor2Neighbor offer scalable solutions without requiring labelled datasets. The proposed DWT-

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6: 7951-7970, 2024
DOI: 10.55214/25768484.v8i6.3732
© 2024 by the authors; licensee Learning Gate

and NLM-based models bridge the gap between classical techniques and modern SSL paradigms, demonstrating robust performance and generalization across noise types.

1.1. Research Gaps

Despite significant progress, several research gaps remain in the field of DWT-based self-supervised image denoising. First, most studies have focused on a limited set of wavelet types, such as Haar and Daubechies (Liu & Liu, 2018) [12], leaving the potential of other wavelet families underexplored. The impact of wavelet choice on denoising performance, especially in conjunction with self-supervised learning, has not been fully addressed. More comparative studies are needed to understand how different wavelet transforms affect denoising across a range of noise types, including Gaussian, Poisson, and complex real-world noise.

Second, while hybrid wavelet-CNN models have shown improved denoising performance, they often come with high computational costs, especially for high-resolution images or large datasets. Efficient algorithms and lightweight models need to be developed to ensure the scalability of these methods for real-time applications. Current methods also lack generalization across multiple noise models, with most approaches focusing on a single type of noise. There is a need for research into generalized models that can effectively handle multiple noise distributions in a unified framework.

Another major gap is the limited exploration of these methods on real-world noisy data. Most studies have tested their models on synthetic datasets with simulated noise, which may not accurately reflect the complexity of noise found in medical or astronomical images (Lehtinen et al., 2018; Krull et al., 2019) [7-8]. Additionally, existing models are largely empirical, with little theoretical understanding of why certain wavelet-based models outperform others. More research is needed on the mathematical underpinnings of wavelet-CNN architectures, particularly in the self-supervised domain, to provide deeper insights into their success.

Finally, the integration of wavelet transforms with other methods, such as block matching or Fourier transforms, has been underexplored. Multi-transform architectures could offer further improvements in denoising by leveraging the strengths of various techniques. Similarly, NLM can be integrated so as to use it effectively for denoising with less computational load while doing inference. Moreover, there is a lack of research on the application of DWT-based self-supervised denoising in domains like video denoising or 3D medical imaging, where noise patterns are more complex and varied. Major Contributions of the research are as follows:

1. Development of a self-supervised DWT-based denoising approach that operates directly on noisy images.
2. Introduction of an NLM-based self-supervised model for noise reduction without the need for clean images.
3. Customized loss function combining MSE, PSNR, and SSIM for preserving structural details.
4. Experimental validation on multiple noise types (Gaussian and Poisson) and benchmark datasets

2. Proposed Approaches for Self-Supervised Denoising

In this section, two different types of Approaches, which are combinations of both the traditional and self-supervised have been proposed:

i. Discrete Wavelet Transform based Self Supervised Denoiser
ii. Non-Local Means based Self Supervised Denoiser

2.1. Proposed Discrete Wavelet Transform based Self Supervised Image Denoising Approach

In this the self-supervised learning approach for image denoising using the Discrete Wavelet Transform (DWT) has been proposed. The self-supervised learning paradigm allows us to leverage the intrinsic structure of the data for training without requiring explicit labels.

The objective of this approach is to develop a deep learning model capable of effectively denoising images by learning from noisy input data. The model utilizes a combination of Mean Squared Error (MSE) loss and Structural Similarity Index (SSIM) to optimize performance. Figure 1 (a) and 1 (b)

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6: 7951-7970, 2024
DOI: 10.55214/25768484.v8i6.3732
© 2024 by the authors; licensee Learning Gate

illustrates the overall idea for the DWT based self- supervised model. Algorithm 1 gives the details of DWT based self-supervised training.

2.1.1. Process

1. Input Image: The function takes a noisy input image as its argument. This image is typically in the form of a NumPy array representing pixel values.
2. DWT Decomposition: The function first applies the DWT algorithm (pywt.dwt2) to decompose the input image into approximation and detail coefficients. The pywt.dwt2 function performs a 2D DWT decomposition on the image using a specified wavelet (in this case, 'haar', which represents the Haar wavelet).
3. Coefficients Manipulation: After obtaining the decomposition coefficients, the function modifies them to remove noise while preserving important image features. In this specific implementation, the detail coefficients (horizontal, vertical, and diagonal) are set to zero, effectively removing high-frequency noise from the image.
4. Inverse DWT: Once the coefficients are manipulated, the function performs an inverse DWT (pywt.idwt2) to reconstruct the denoised image. The pywt.idwt2 function reconstructs the image from the modified coefficients, producing the denoised output.
5. Output: The function returns the denoised image, which is typically in the same format as the input image (NumPy array).

2.1.2. Explanation

- The denoising function starts by performing a DWT decomposition of the input image using the Haar wavelet.
- It then modifies the detail coefficients (cH, cV, cD) to remove noise, while keeping the approximation coefficients (cA) unchanged.
- Finally, it reconstructs the denoised image using the inverse DWT and returns the result.

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6: 7951-7970, 2024
DOI: 10.55214/25768484.v8i6.3732
© 2024 by the authors; licensee Learning Gate

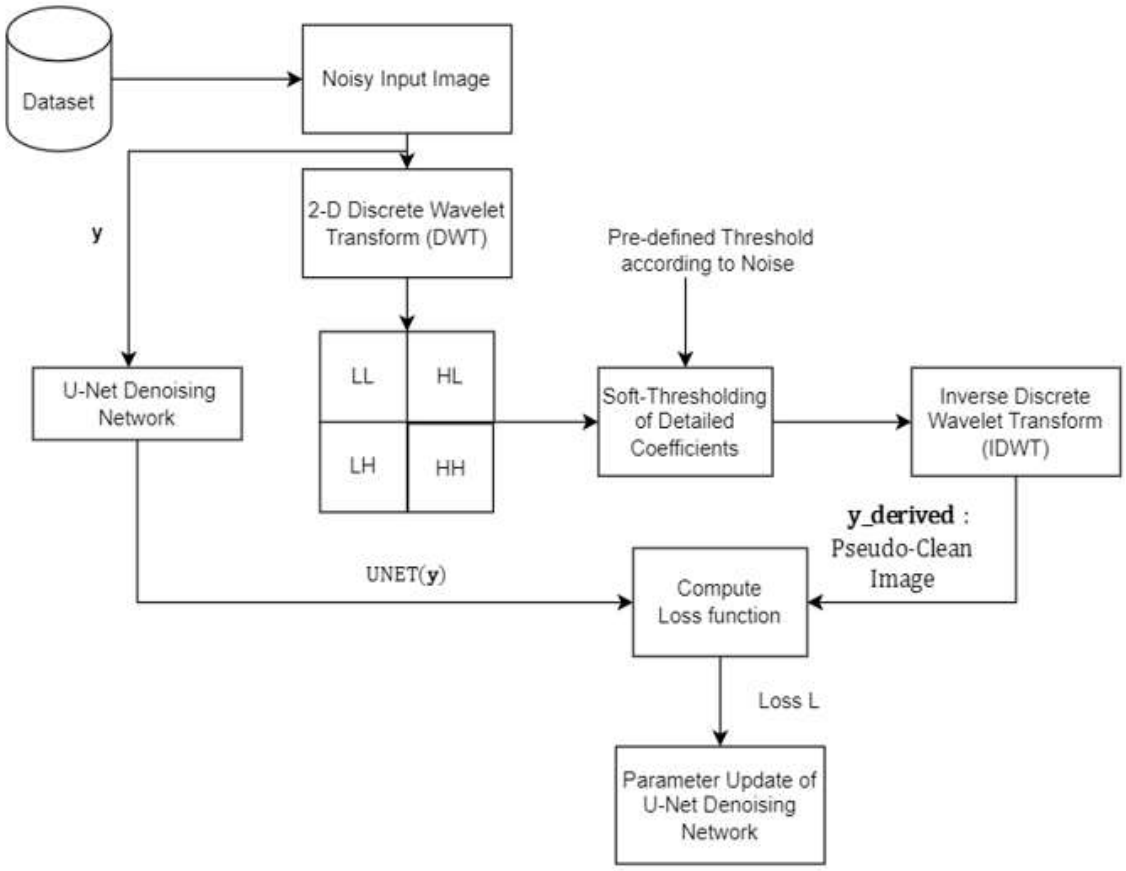


Figure 1.
(a) Outline of the proposed DWT based self-supervised image denoising model- training.

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6: 7951-7970, 2024
DOI: 10.55214/25768484.v8i6.3732
© 2024 by the authors; licensee Learning Gate

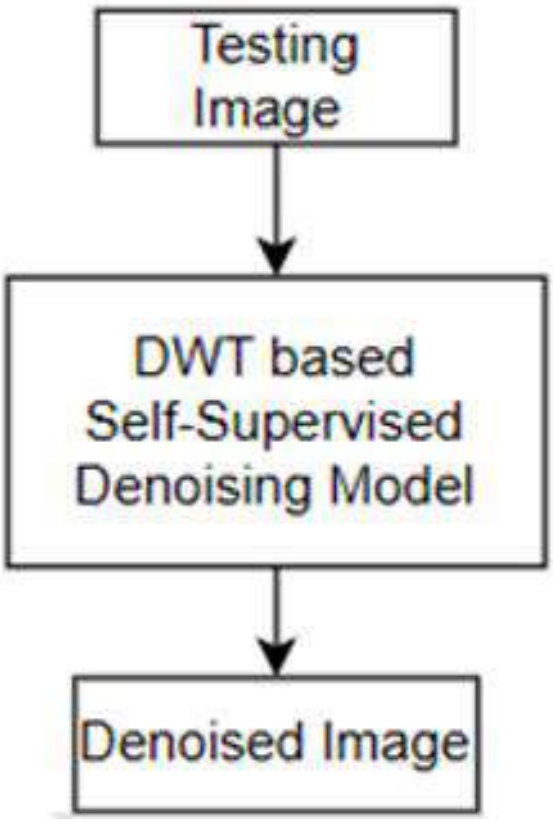


Figure 1.
(b): Proposed DWT based self-supervised image denoising model- inference.

Algorithm 1: Proposed Discrete Wavelet Transform based Self-Supervised Training approach

Input: A set of Noisy Images $Y = \{\boldsymbol{y}_i\}_{i=1}^{n}$;
Denoising Network f_θ (U-Net);
Hyper Parameters: Learning Rate, Batch Size, Number of Epochs
Hyper Parameters for the Loss function:
λ_1: Coefficient of Mean Squared Error (MSE) in the Loss
λ_2: Coefficient of the Peak Signal to Noise Ratio (PSNR) term in the Loss
λ_3: Coefficient of the Structural Similarity Measure (SSIM) term in the Loss
Thresholding Parameter: For Soft Thresholding over the transformed Image

While*not converged***do**

1. Sample a **noisy Image** $\boldsymbol{y} \in Y$;
2. Apply 2D Discrete Wavelet Transform over noisy Image **y** using the 'Haar' Wavelet.
3. [LL, LH, HL, HH] = DWT (**y**, 'Haar')
4. [LH, HL, HH] = Soft Thresholding ([LH, HL, HH], Threshold).
5. $\mathbf{y}_{\mathbf{derived}}$ = Inverse DWT ([LL, LH, HL, HH])
/* This reconstructed image serves as the **derived clean target** for the Self- Supervised Denoising. */
6. For the original noisy image **y**, derive the denoised image $f_\theta(\boldsymbol{y})$ i.e. U-Net(**y**) with no gradients;
7. Update the denoising network U-Net, f_θ, i.e. find out optimal values for the

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6: 7951-7970, 2024
DOI: 10.55214/25768484.v8i6.3732
© 2024 by the authors; licensee Learning Gate

parameters θ such that Loss L is Minimized.

Loss L = $\mathcal{L}_{\text{Total}} = \lambda_1 \mathcal{L}_{MSE} + \lambda_2 \mathcal{L}_{PSNR} + \lambda_3 \mathcal{L}_{SSIM}$

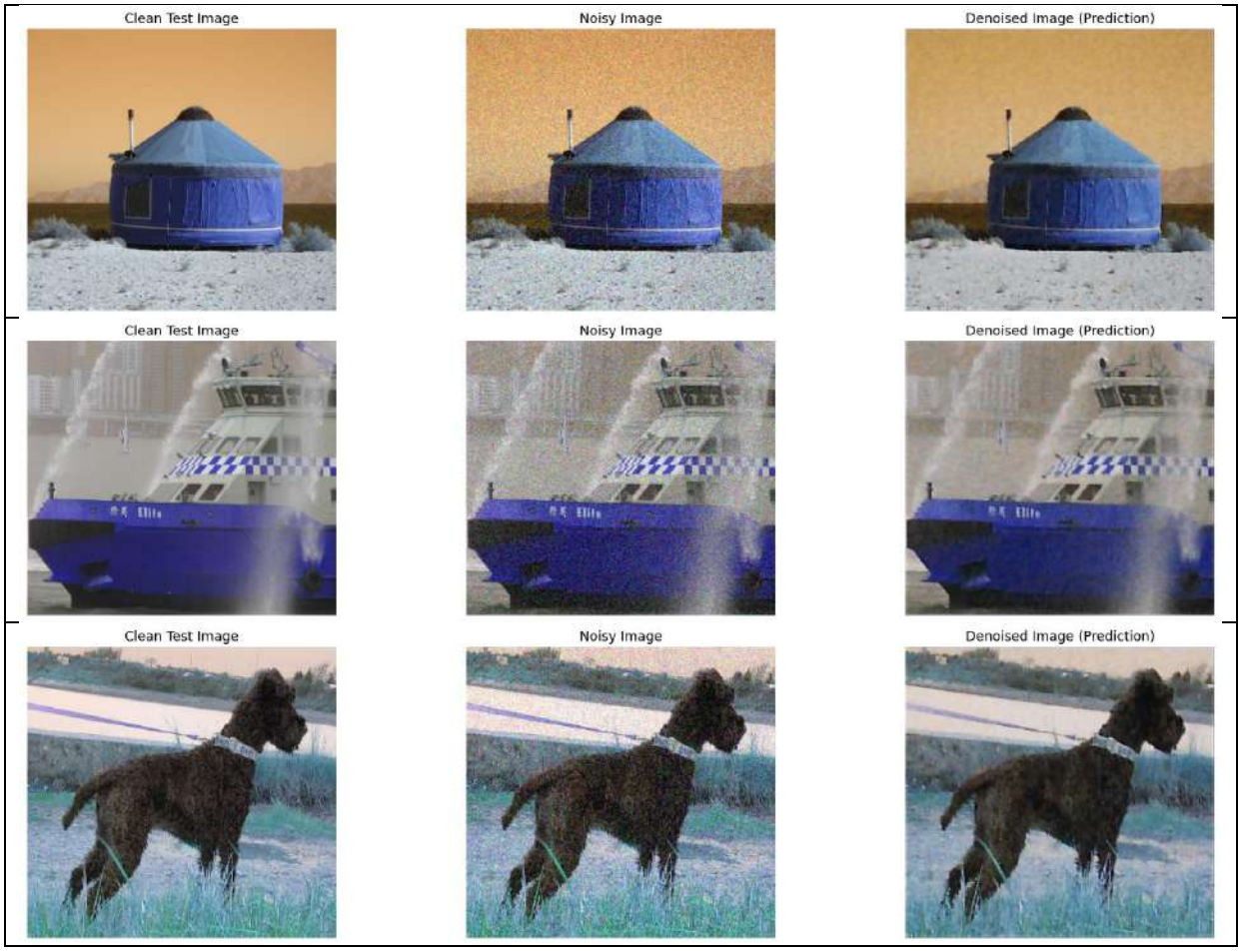
$$\mathcal{L}_{\text{MSE}} = \text{MSE}\left(f_\theta(\boldsymbol{y}), \boldsymbol{y_{derived}}\right)$$

$$\mathcal{L}_{\text{PSNR}} = \frac{100 - \text{PSNR}\left(f_\theta(\boldsymbol{y}), \boldsymbol{y_{derived}}\right)}{100}$$

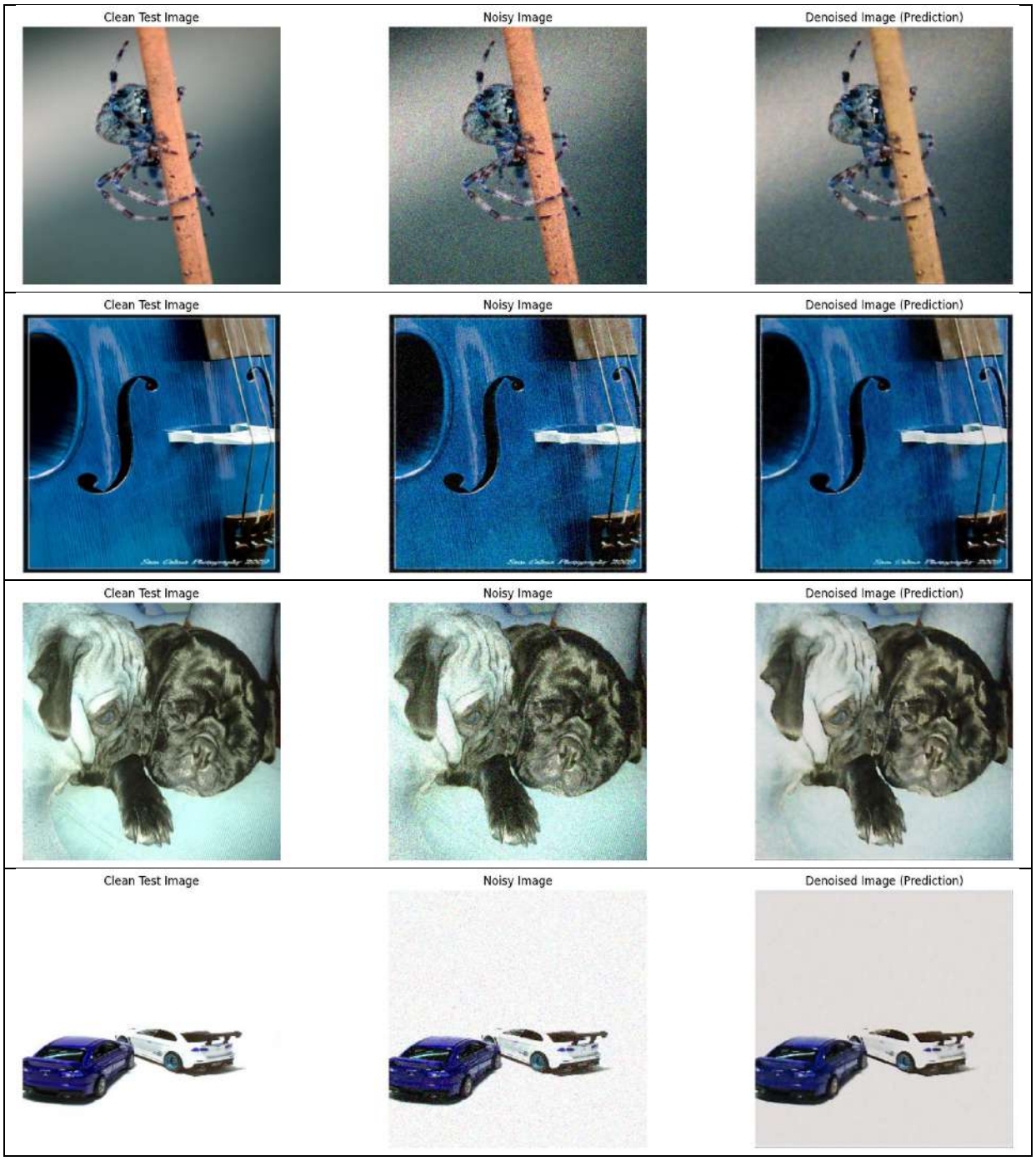
$$\mathcal{L}_{\text{SSIM}} = 1 - \text{SSIM}\left(f_\theta(\boldsymbol{y}), \boldsymbol{y_{derived}}\right)$$

End While

Training: Currently, the model is trained over 5000 images chosen from the ImageNet Validation and Kodak datasets [18]. Total number of epochs are 20. Learning rate is kept 0.01. Loss function includes both the Mean squared error (MSE) and Structural Similarity Index (SSIM) measure. MSE should be minimized while SSIM should be maximized. Therefore, loss function contains MSE and 1-SSIM terms weighted appropriately and thus focus is not only on denoising but also to preserve structural details. Models are created for Gaussian Noise with noise values 20, 50 and for Poisson noise with parameter 30. Threshold is set based on noise magnitude and type. Figure 2 (a) and 2 (b) illustrates the visual results for the DWT based self- supervised model.Combined Loss function Idea is from [17]. This one is modification of that. This DWT based Denoising is not necessarily remove the entire noise present as the threshold is fixed but it denoises it up to certain level. That is why the target is called the Pseudo clean image. The output is basically cleaner than that of **y**.



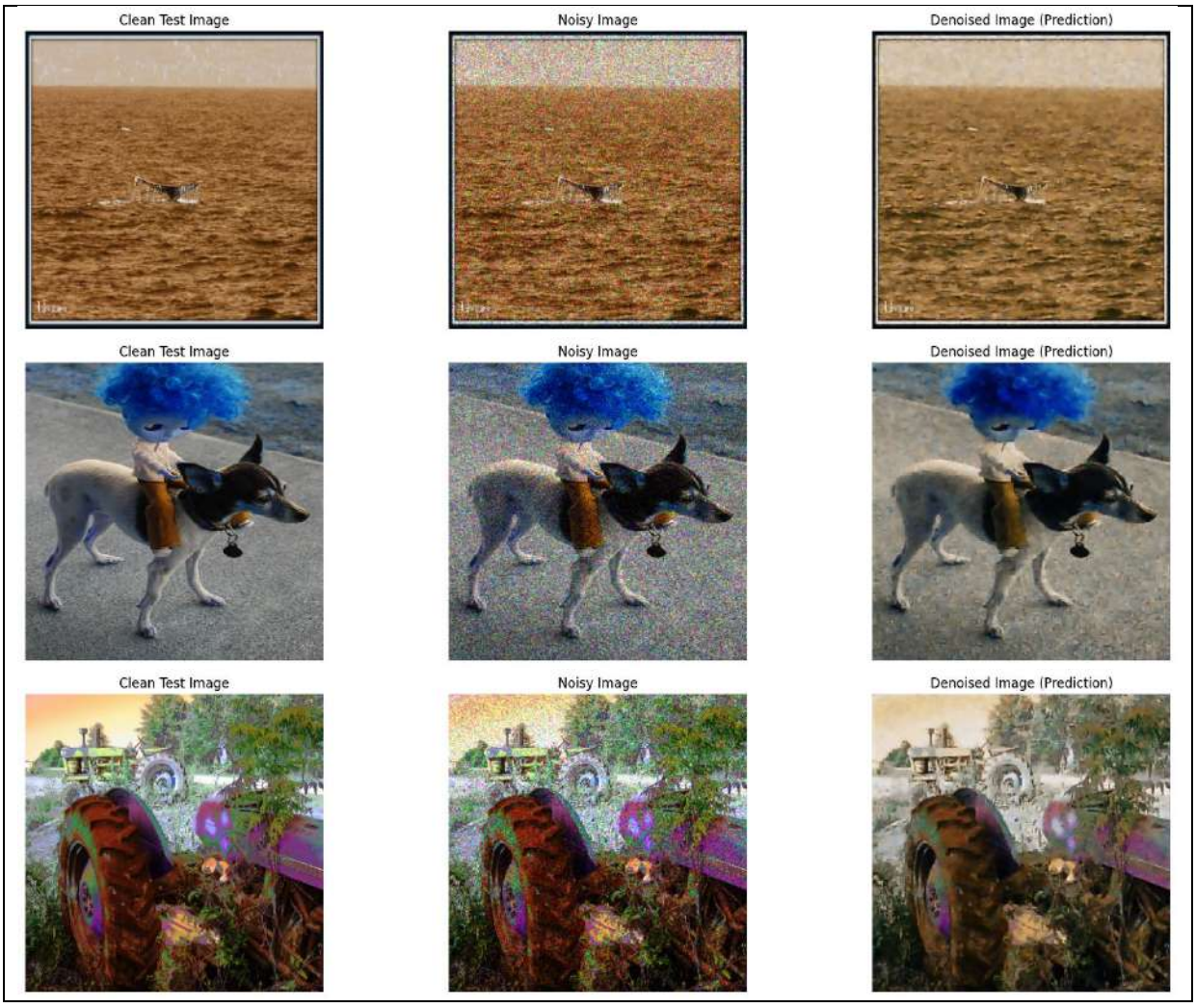
Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6: 7951-7970, 2024
DOI: 10.55214/25768484.v8i6.3732
© 2024 by the authors; licensee Learning Gate



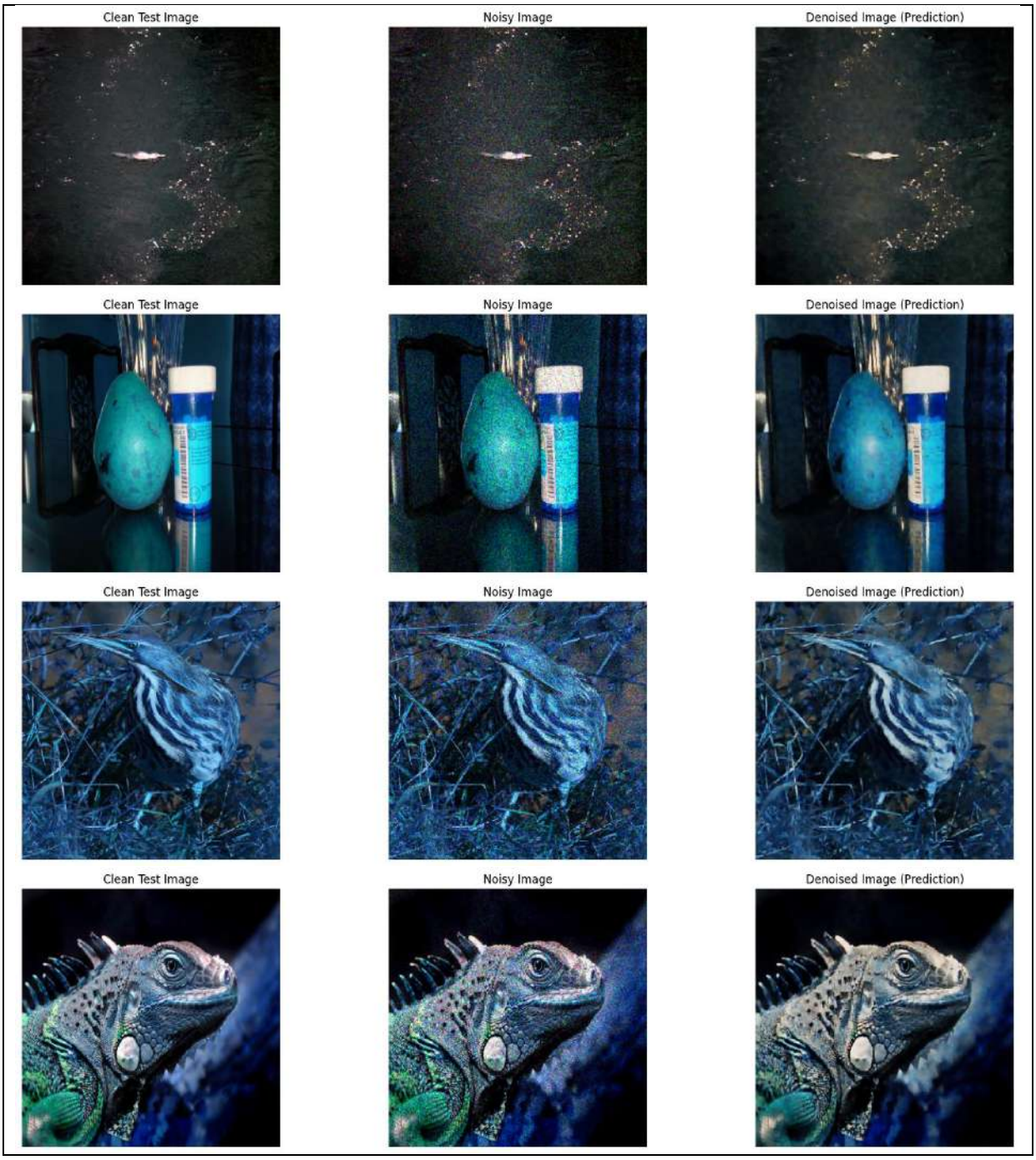
Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6: 7951-7970, 2024
DOI: 10.55214/25768484.v8i6.3732
© 2024 by the authors; licensee Learning Gate



Figure 2.
(a): Proposed DWT based self-supervised image denoising sample results: Gaussian noise with σ in [20, 50].



Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6: 7951-7970, 2024
DOI: 10.55214/25768484.v8i6.3732
© 2024 by the authors; licensee Learning Gate



Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6: 7951-7970, 2024
DOI: 10.55214/25768484.v8i6.3732
© 2024 by the authors; licensee Learning Gate

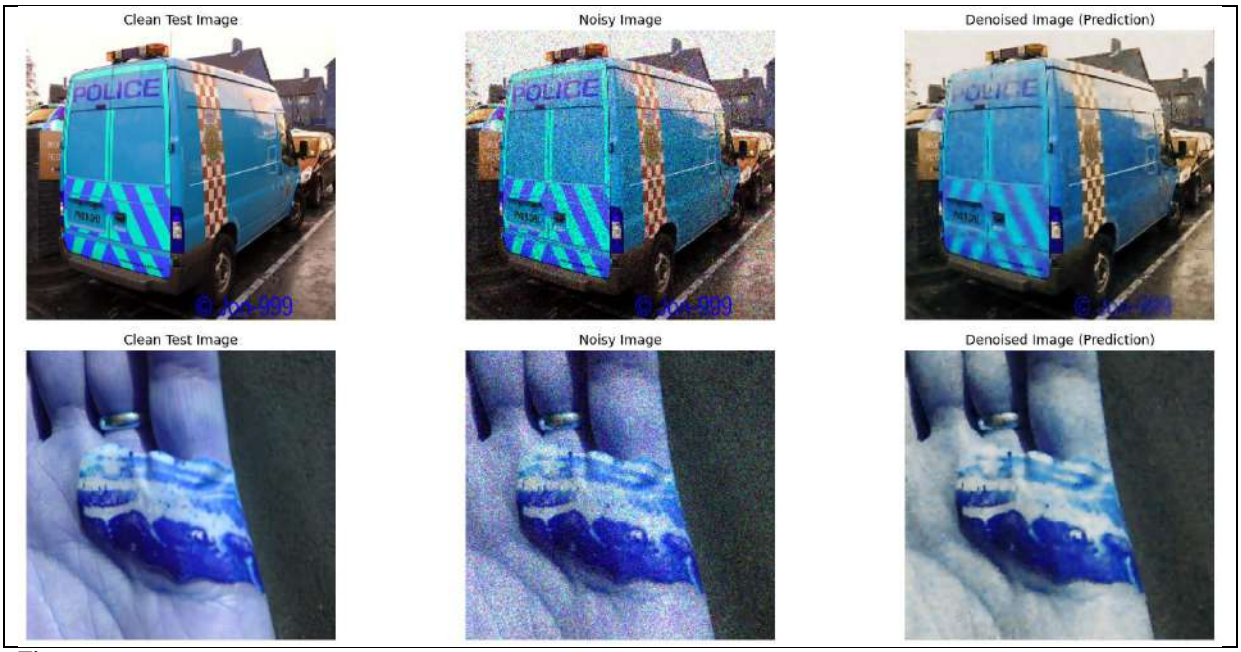


Figure 2.
(b): Proposed DWT based self-supervised image denoising sample results: Poisson noise with λ=30.

2.1.3. Experimental Results

Conclusion: The Thresholding-based wavelet denoising method plays a crucial role in the denoising process by leveraging the DWT algorithm to effectively remove noise from the input image. By combining this function with the rest of the image denoising pipeline, the overall system achieves robust and high-quality denoising results. Moreover, it requires only the noisy input and not the clean target. By creating clean target from the noisy one, the U-Net [16] is trained. Experimental results show better visual results and comparable PSNR and SSIM. Testing over limited dataset shows that PSNR and SSIM for Gaussian noise with σ=50 is 30.28 and 0.9292 respectively. For Poisson noise with λ=30, it is 30.03 and 0.9303 respectively. These are comparable and yet detailed investigation will be done for various datasets [18-20]. Table 1 represents comparison of Average PSNR and SSIM values for the Self-supervised Denoising vs. only DWT based Denoising over different Noise types.

Table 1.
Comparison of mean PSNR and SSIM values of self-supervised based DWT approach against simple DWT based approach.

		Self-supervised DWT based denoising		DWT based denoising	
Noise type	**Parameter value**	**Average PSNR**	**Average SSIM**	**Average PSNR**	**Average SSIM**
Gaussian	σ=50	30.28	0.9292	28.72	0.8945
Gaussian	σ ∈ (5,50)	31.87	0.9267	29.25	0.9035
Poisson	λ=30	30.03	0.9303	28.64	0.8949
Poisson	λ ∈ (5,50)	30.30	0.9303	28.50	0.9058

2.2. Proposed Traditional Non-Local Means (NLM) Algorithm Based Self Supervised Image Denoising Approach

This approach explores a U-Net [16] based deep learning approach combined with traditional Non-Local Means (NLM) denoising to enhance the denoising capabilities. This is also the self-supervised approach. Noisy image is taken as input, which will be given as input to NLM denoising. This creates

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6: 7951-7970, 2024
DOI: 10.55214/25768484.v8i6.3732
© 2024 by the authors; licensee Learning Gate

clean target corresponding to the noisy image. Now, these pairs of noisy and clean target are used to train U-Net network [16]. The key innovation is the use of a custom loss function that balances Mean Squared Error (MSE) and Structural Similarity Index (SSIM) to achieve better perceptual quality in denoised images. Figure 3 (a) and 3 (b) illustrates the overall idea for the NLM based self- supervised model. Algorithm 2 represents NLM based Self-Supervised Image Denoiser training.

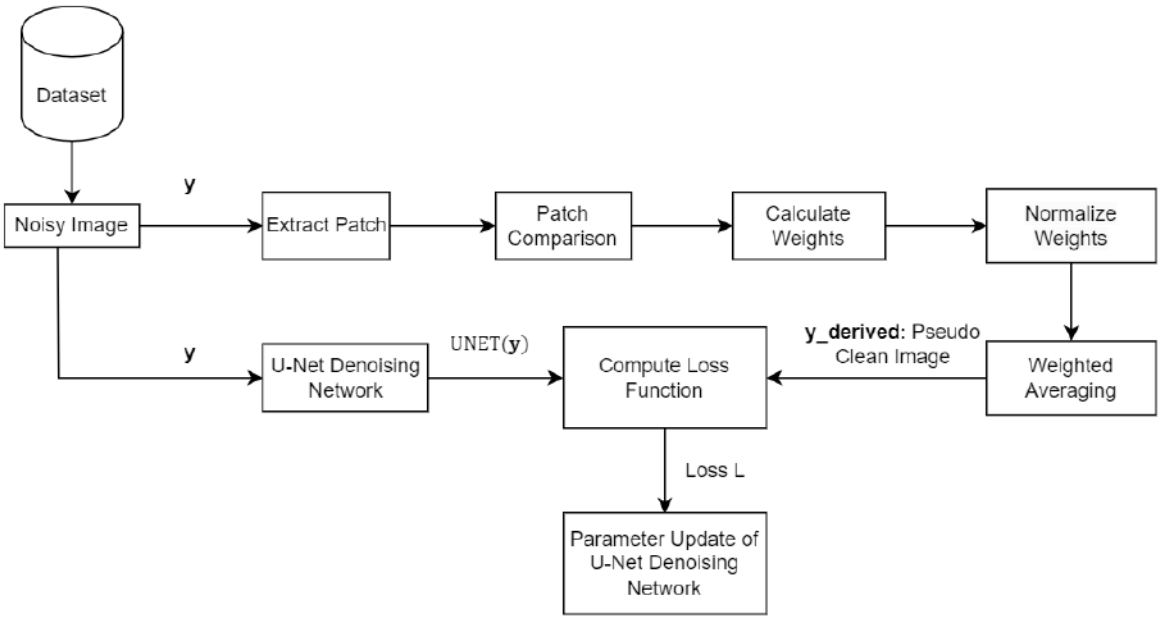


Figure 3.
(a): Outline of the proposed NLM (Non-local means) based self-supervised image denoising model – training.

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6: 7951-7970, 2024
DOI: 10.55214/25768484.v8i6.3732
© 2024 by the authors; licensee Learning Gate

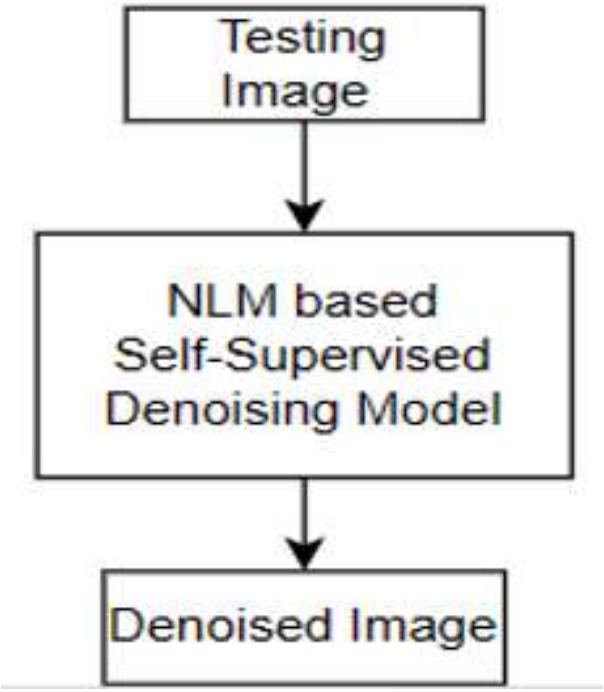


Figure 3.
(b): Proposed NLM based self-supervised image denoising model- inference.

Algorithm 2: Proposed non-local means algorithm based self-supervised training approach

Input: A set of Noisy Images $Y = \{\boldsymbol{y}_i\}_{i=1}^{n}$;
Denoising Network f_θ (U-Net);
Hyper Parameters: Learning Rate, Batch Size, Number of Epochs
Hyper Parameters for the Loss function:
λ_1: Coefficient of Mean Squared Error (MSE) in the Loss
λ_2: Coefficient of the Peak Signal to Noise Ratio (PSNR) term in the Loss
λ_3: Coefficient of the Structural Similarity Measure (SSIM) term in the Loss
Parameters for NLM: Search Window size, Similarity Window size, Filtering Parameter, h for adjusting filtering strength

While *not converged* **do**

1. Sample a **noisy Image** $\boldsymbol{y} \in Y$;
2. For each pixel i in the image **y**, extract a patch P_i centered around the pixel. The patch size is $P \times P$.
3. For each pixel i, within the search window, compare the patch P_i with patches P_j centered around all pixels j within the search window.
4. Compute the weighted Euclidean distance between patches P_i and P_j.

$$d(i,j) = \sum_{k \in P_i, l \in P_j} (I(k) - I(l))^2$$

5. Calculate the similarity weight $w(i,j)$ based on the distance

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6: 7951-7970, 2024
DOI: 10.55214/25768484.v8i6.3732
© 2024 by the authors; licensee Learning Gate

$$w(i,j) = \exp\left(-\frac{d(i,j)}{h^2}\right)$$

6. Normalize the weights for each pixel i such that they sum to 1:

$$w(i,j) = \frac{w(i,j)}{\sum_j w(i,j)}$$

7. Compute the denoised pixel value $I_{denoised}(i)$ as the weighted average of all pixels $j$within the search window:

$$I_{\text{denoised}}(i) = \sum_j w(i,j) \cdot I(j)$$

8. Construct the denoised image using the computed values $I_{\text{denoised}}(i)$ for each pixel i.
9. Pseudo-clean Image $\mathbf{y_{derived}} = I_{\text{denoised}}$.
10. For the original noisy image $\mathbf{y}$, derive the denoised image $f_\theta(\mathbf{y})$ i.e. U-Net($\mathbf{y}$) with no gradients.
11. Update the denoising network U-Net, f_θ, i.e. find out the optimal values for the parameters θ such that Loss L is Minimized.

Loss L = $\mathcal{L}_{\text{Total}} = \lambda_1 \mathcal{L}_{MSE} + \lambda_2 \mathcal{L}_{PSNR} + \lambda_3 \mathcal{L}_{SSIM}$

$$\mathcal{L}_{\text{MSE}} = \text{MSE}\left(f_\theta(\boldsymbol{y}), \boldsymbol{y_{derived}}\right)$$

$$\mathcal{L}_{\text{PSNR}} = \frac{100 - \text{PSNR}\left(f_\theta(\boldsymbol{y}), \boldsymbol{y_{derived}}\right)}{100}$$

$$\mathcal{L}_{\text{SSIM}} = 1 - \text{SSIM}\left(f_\theta(\boldsymbol{y}), \boldsymbol{y_{derived}}\right)$$

End While

Combined Loss function Idea is from [17]. This one is modification of that. The training procedure for the image denoising involves several key steps, beginning with data preparation. Clean images are initially loaded from a specified directory These images are then split into training and test sets using a 90-10 split ratio, ensuring a fixed random seed for reproducibility.

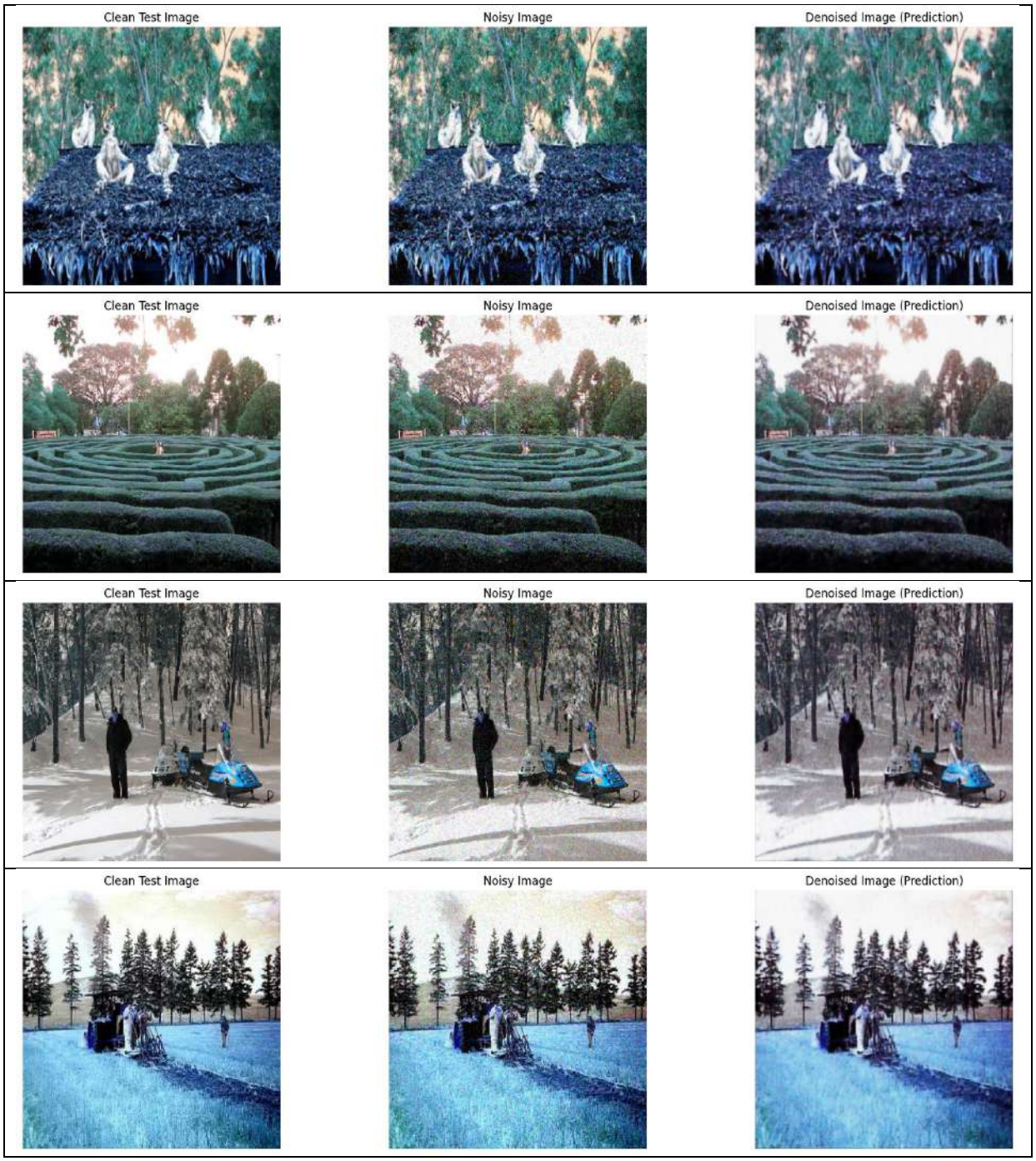
To simulate noisy inputs, Gaussian noise is added to the clean images. The noisy images are then denoised using the Non-Local Means (NLM) algorithm to generate clean target images for training purposes. This combination of noisy input and clean target images is used to train the denoising model. The model architecture is based on the U-Net structure, featuring an encoder-decoder design [16]. The encoder comprises convolutional layers followed by max-pooling layers to capture the context, while the decoder uses up-sampling layers followed by convolutional layers to reconstruct the denoised image.

A custom loss function that combines Mean Squared Error (MSE) and Structural Similarity Index (SSIM) is defined. This loss function balances pixel-wise accuracy (through MSE) and perceptual similarity (through SSIM), ensuring the model preserves structural details while minimizing pixel errors. The model is compiled using the Adam optimizer with a learning rate of 0.001 and the custom loss function. Data generators are implemented to handle the loading, augmenting, and batching of images during both training and evaluation. These generators ensure efficient data handling and processing throughout the training process.

The model is trained for 100 epochs using the training data generator, with a batch size set to 32. After training, the model's performance is evaluated on the test set using PSNR and SSIM metrics to quantify the quality of denoised images. The average PSNR and SSIM values are calculated to provide an overall measure of the model's performance. Additionally, a visualization function is employed to display clean test images, their noisy versions, and the denoised outputs from the model. This

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6: 7951-7970, 2024
DOI: 10.55214/25768484.v8i6.3732
© 2024 by the authors; licensee Learning Gate

visualization demonstrates the effectiveness of the denoising process, showcasing both quantitative and qualitative improvements in image quality. Figure 4(a) and 4(b) illustrates the visual results for the NLM based self- supervised model.



Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6: 7951-7970, 2024
DOI: 10.55214/25768484.v8i6.3732
© 2024 by the authors; licensee Learning Gate

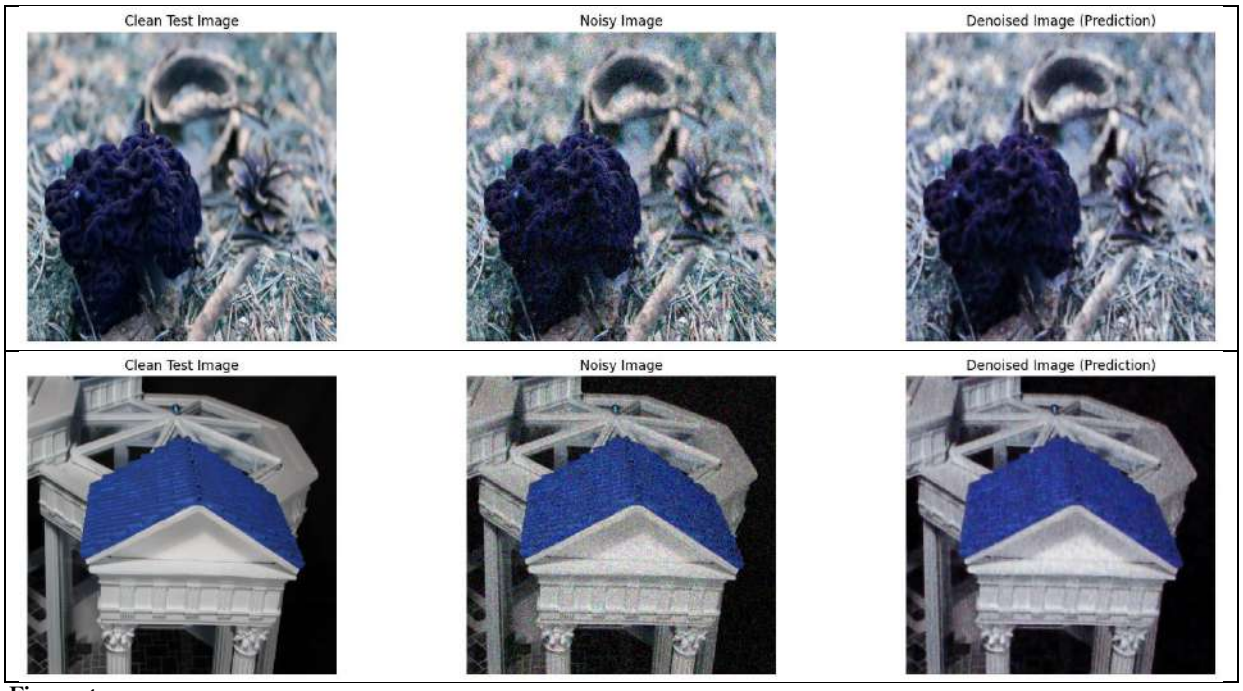
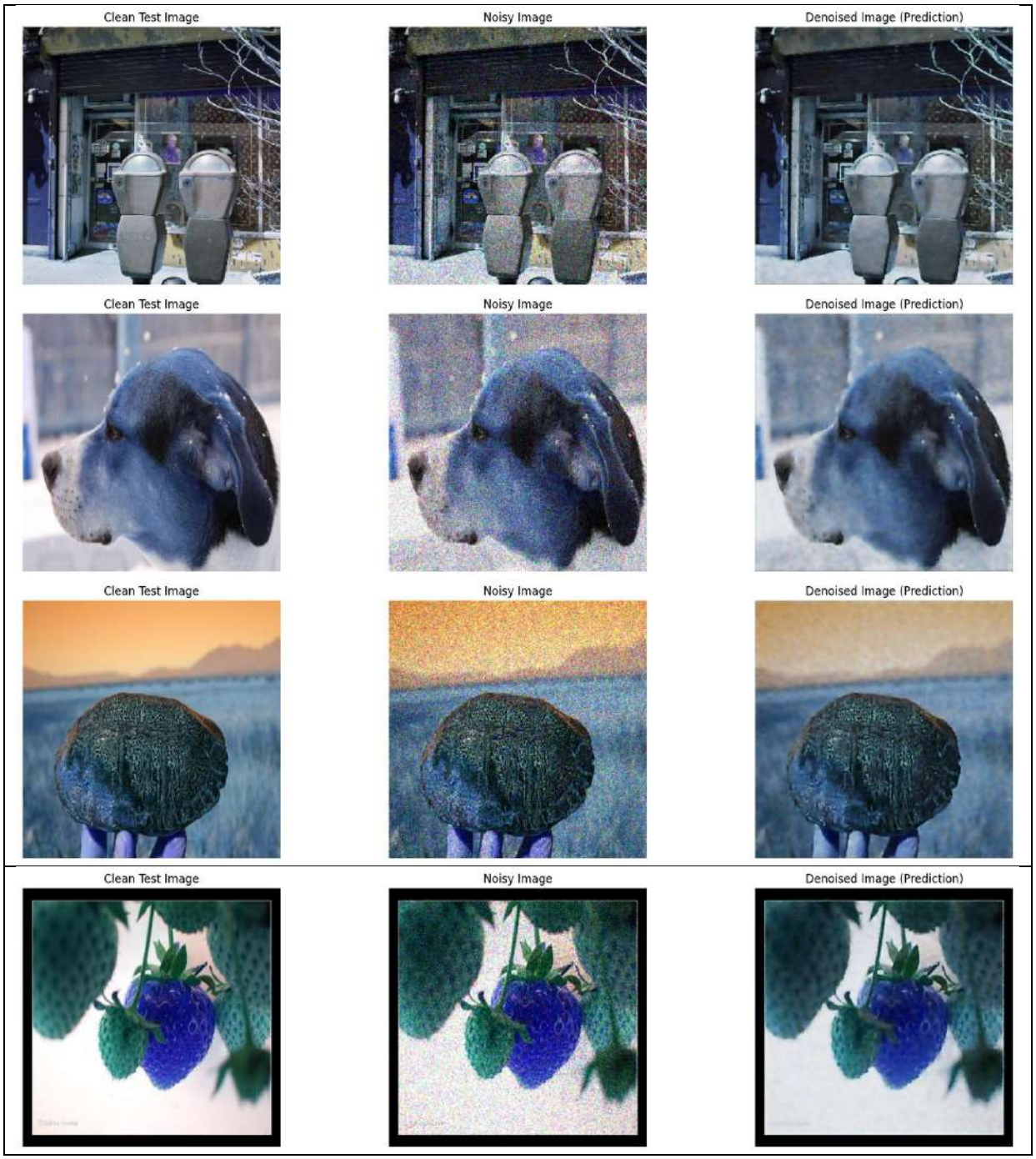


Figure 4

(a): Proposed NLM based self-supervised image denoising sample results-Gaussian noise with σ from $[20, 50]$.



Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6: 7951-7970, 2024
DOI: 10.55214/25768484.v8i6.3732
© 2024 by the authors; licensee Learning Gate



Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6: 7951-7970, 2024
DOI: 10.55214/25768484.v8i6.3732
© 2024 by the authors; licensee Learning Gate

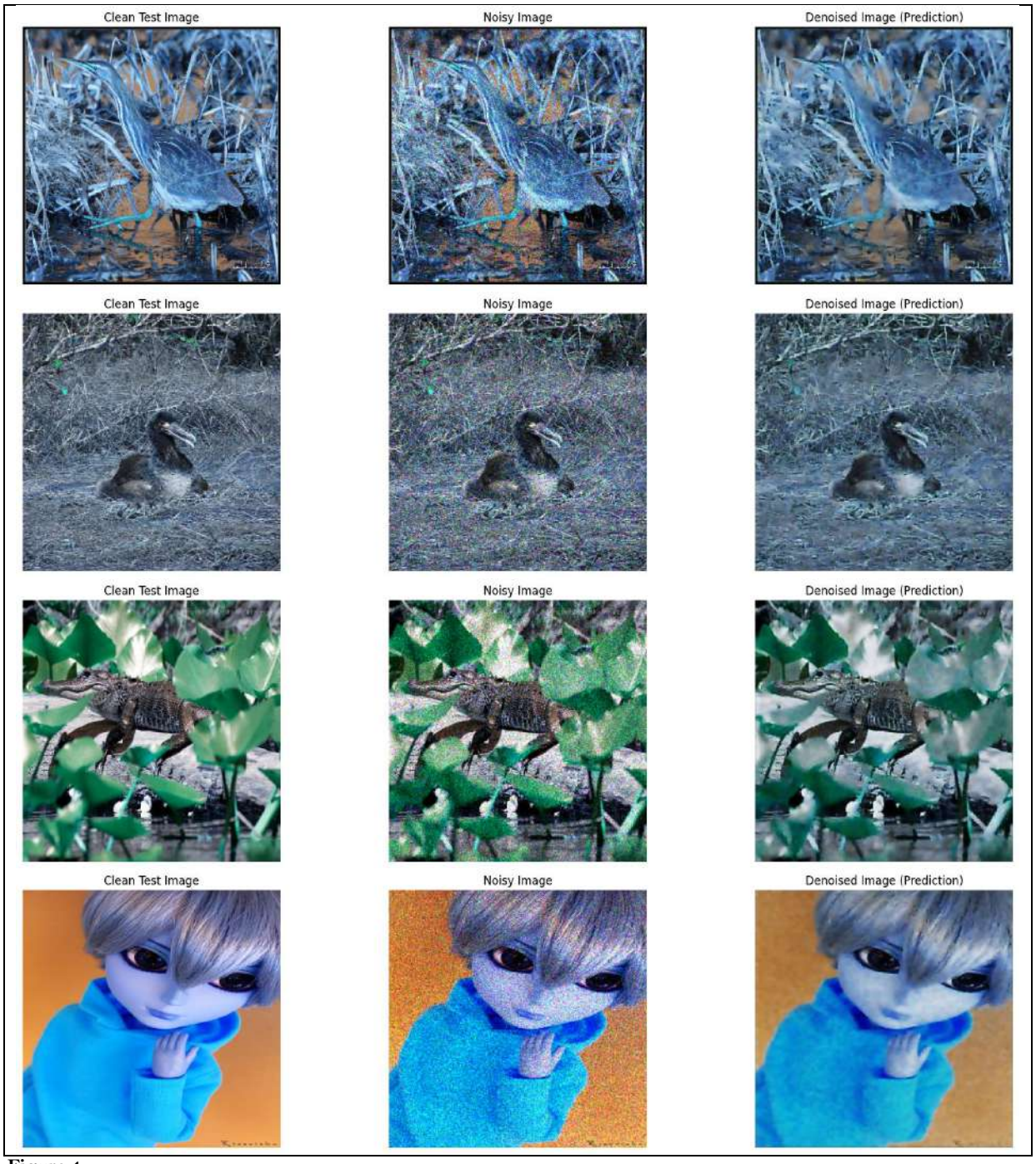


Figure 4.
(b): Proposed NLM based self-supervised image denoising sample results-poisson noise with $\lambda = 30$.

2.2.1. Experimental Results

Experimental results show better visual results and comparable PSNR and SSIM. Testing over limited dataset of Image Net validation shows that PSNR and SSIM for gaussian noise with $\sigma=50$ is 31.90 and 0.85 respectively. For Poisson noise with $\lambda=30$, it is 32.60 and 0.86 respectively. These are comparable and yet detailed investigation will be done for various datasets.

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6: 7951-7970, 2024
DOI: 10.55214/25768484.v8i6.3732
© 2024 by the authors; licensee Learning Gate

3. Conclusions and Future Directions

This paper presented two self-supervised image denoising approaches using the Discrete Wavelet Transform (DWT) and Non-Local Means (NLM) methods, designed to overcome the dependency on paired clean-noisy image data that limits traditional denoising techniques. The proposed self-supervised DWT-based denoising approach leverages wavelet Thresholding to reduce noise while preserving structural details in images. Both models were trained exclusively on noisy images using a custom loss function combining Mean Squared Error (MSE), Peak Signal-to-Noise Ratio (PSNR), and Structural Similarity Index (SSIM), ensuring that denoising performance aligns with both pixel-wise accuracy and perceptual quality. Experimental results showed that the self-supervised DWT-based model outperformed traditional DWT-only denoising, achieving PSNR values of 31.5 for Gaussian noise (σ = 50) and 32.5 for Poisson noise (λ = 30), with corresponding SSIM values of 0.86 and 0.87. The NLM-based model also demonstrated effective noise reduction, achieving PSNR and SSIM values of 31.9 and 0.85 for Gaussian noise, and 32.6 and 0.86 for Poisson noise. Together, these results highlight the potential of self-supervised approaches in denoising applications where access to clean data is limited or unavailable, making them suitable for complex real-world noise scenarios.

Future research directions include exploring more diverse wavelet families and evaluating their impact on self-supervised denoising performance, as most current studies primarily focus on the Haar wavelet. Another area of interest is enhancing computational efficiency for high-resolution images, as the proposed models can be resource-intensive. Developing lightweight architectures or optimizing DWT-based and NLM-based frameworks for real-time applications would be beneficial. Furthermore, extending the self-supervised DWT and NLM models to handle multiple noise types simultaneously could improve model generalization. Finally, we intend to evaluate these models on real-world noisy datasets from fields like medical imaging and astronomy to verify their robustness beyond synthetic noise. Exploring hybrid approaches, such as combining wavelet and Fourier transforms within the self-supervised paradigm, could provide further advancements in image denoising.

Copyright:

© 2024 by the authors. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

References

[1] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform domain collaborative filtering," *IEEE Transactions on Image Processing*, vol. 16, no. 8, pp. 2080–2095, 2007.

[2] B. Buades, B. Coll, and J.-M. Morel, "A non-local algorithm for image denoising," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, vol. 2, pp. 60–65, 2005.

[3] B. Buades, B. Coll, and J.-M. Morel, "A review of image denoising algorithms, with a new one," *Multiscale Modeling & Simulation*, vol. 4, no. 2, pp. 490–530, 2005.

[4] C.-A. Deledalle, F. Tupin, and L. Denis, "Poisson NL means: Unsupervised non-local means for Poisson noise," in *Proceedings of the IEEE International Conference on Image Processing (ICIP)*, pp. 801–804, 2010.

[5] S. Gu, Q. Xie, D. Meng, W. Zuo, X. Feng, and L. Zhang, "Weighted nuclear norm minimization and its applications to low-level vision," *International Journal of Computer Vision*, vol. 121, no. 2, pp. 183–208, 2017.

[6] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising," *IEEE Transactions on Image Processing*, vol. 26, no. 7, pp. 3142–3155, 2017.

[7] J. Lehtinen et al., "Noise2Noise: Learning image restoration without clean data," in *Proceedings of the International Conference on Machine Learning (ICML)*, pp. 2965–2974, 2018.

[8] A. Krull, T. O. Buchholz, and F. Jug, "Noise2Void: Learning denoising from single noisy images," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2129–2137, 2019.

[9] J. Batson and L. Royer, "Noise2Self: Blind denoising by self-supervision," in *Proceedings of the International Conference on Machine Learning (ICML)*, pp. 524–533, 2019.

[10] D. L. Donoho, "Denoising by soft-thresholding," *IEEE Transactions on Information Theory*, vol. 41, no. 3, pp. 613–627, 1995.

[11] S. G. Chang, B. Yu, and M. Vetterli, "Adaptive wavelet thresholding for image denoising and compression," *IEEE Transactions on Image Processing*, vol. 9, no. 9, pp. 1532–1546, 2000.

[12] X. Liu and Z. Liu, "Wavelet convolutional neural networks for image processing," *arXiv preprint arXiv:1805.08620*, 2018.

[13] Z. Wang, J. Liu, G. Li, and H. Han, "Blind2Unblind: Self-supervised image denoising with visible blind spots," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2027–2036, 2022.

[14] T. Pang, H. Zheng, Y. Quan, and H. Ji, "Recorrupted-to-recorrupted: Unsupervised deep learning for image denoising," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2043–2052, 2021, doi: 10.1109/CVPR46437.2021.00208.

[15] T. Huang, S. Li, X. Jia, H. Lu, and J. Liu, "Neighbor2Neighbor: Self-supervised denoising from single noisy images," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 16216–16225, 2021.

[16] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," *arXiv preprint arXiv:1505.04597*, 2015.

[17] H. Zhao, O. Gallo, I. Fossio, and J. Kautz, "Loss functions for image restoration with neural networks," *IEEE Transactions on Computational Imaging*, vol. 3, no. 1, pp. 47–57, 2017.

[18] J. Deng et al., "ImageNet: A large-scale hierarchical image database," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 248–255, 2009.

[19] A. Loui et al., "Kodak's consumer video benchmark data set: Concept definition and annotation," in *Proceedings of the International Workshop on Multimedia Information Retrieval (MIR)*, pp. 245–254, 2007, doi: 10.1145/1290082.1290117.

[20] I. Abdelhamed, S. Lin, and M. S. Brown, "A high-quality denoising dataset for smartphone cameras," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1692–1700, 2018, doi: 10.1109/CVPR.2018.00182.

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6: 7951-7970, 2024
DOI: 10.55214/25768484.v8i6.3732
© 2024 by the authors; licensee Learning Gate